基于通道的内源性注意对声音诱发闪光错觉的影响*

张　明[1]　唐晓雨[2]　于　薇[3]　宁　波[1]　王智楠[1]　王爱君[1]

([1]苏州大学心理学系，心理与行为科学研究中心，苏州 215123)

([2]辽宁师范大学心理学院，儿童青少年健康人格评定与培养协同创新中心，大连 116029)

([3]长春中医药大学招生与就业处，长春 130117)

摘　要　声音诱发闪光错觉是指当视觉闪光伴随不相等数量的听觉声音在 100 ms 内相继或同时呈现时，个体会错觉性地知觉到视觉闪光的个数与听觉声音的数量相等。研究采用经典的声音诱发闪光错觉的范式，将注意以内源性的方式指向视觉通道或者听觉通道，考察了基于通道的内源性注意对声音诱发闪光错觉的影响。结果发现，当注意内源性地指向视觉通道时，相对于基线条件，裂变错觉的量显著减小；当注意内源性地指向听觉通道时，相对于基线条件，裂变错觉的量存在增大的趋势。说明了基于通道的内源性注意可以影响声音诱发闪光错觉中的裂变错觉，但不影响融合错觉。

关键词　内源性注意；声音诱发闪光错觉；听觉主导效应；视觉和听觉通道

分类号　B842

1　引言

声音诱发闪光错觉(sound-induced flash illusion)是一种听觉占主导性的多感觉整合现象，即当视觉闪光伴随不相等数量的听觉声音在 100 ms 内相继或同时呈现时，个体会错觉性地知觉到视觉闪光的个数与听觉声音的数量相等。这个现象说明了听觉信息优先于视觉信息的加工，即听觉信息可以主导视觉信息(Abadi & Murphy, 2014; Cecere, Rees, & Romei, 2015; van Erp, Philippi, & Werkhoven, 2013; Kamke, Vieth, Cottrell, & Mattingley, 2012; Kumpik, Roberts, King, & Bizley, 2014; Shams, Kamitani, & Shimojo, 2000, 2002; Whittingham, Mcdonald, & Clifford, 2014)。声音诱发闪光错觉存在两种形式，一种是当一个视觉闪光伴随两个听觉声音时，单个视觉闪光会被错误知觉为两个视觉闪光，也被称为裂变错觉(fission illusion) (Shams et al., 2000, 2002)；另一种是当两个闪光伴随一个听觉声音时，两个视觉闪光会被错误知觉为一个视觉闪光，也被称为融合错觉(fusion illusion) (Andersen, Tiippana, & Sams, 2004)。研究表明，相对于裂变错觉而言,融合错觉的效应较弱(Shams et al., 2000; Wozny, Beierholm, & Shams, 2008)。

收稿日期: 2017-12-18

*国家自然科学基金(31371025, 31600882, 31700939)、江苏省基础研究计划(BK20170333)、教育部人文社科基金(17YJC190024)、江苏省高校哲社科基金(2017SJB1344)和中国博士后基金(2017M611888)资助。

唐晓雨和张明同为第一作者

通信作者：张明, E-mail: psyzm@suda.edu.cn；王爱君, E-mail: ajwang@suda.edu.cn

自从 Shams 等人(2000)发现声音诱发闪光错觉现象后，许多研究者开始研究影响声音诱发闪光错觉现象的因素，研究者们认为,来自自上而下(top-down)的注意可以影响声音诱发闪光错觉现象。以往研究发现，声音诱发闪光错觉在第二个听觉声音呈现 30~60 ms 后产生了对视觉皮层活动幅度的早期调节作用(Mishra, Martínez, Sejnowski, & Hillyard, 2007)。Mishra, Mishra, Martínez 和 Hillyard (2010)的研究进一步考察了刺激位置(高视野和低视野)与注意(注意和不注意)的交互作用对声音诱发闪光错觉的影响。结果表明，PD120/110 成分(在低视野)位于腹侧枕、颞外纹视觉皮层，而随后的 PD180 和 ND250/240 成分(在低视野)位于颞上回(枕颞皮层和颞上回被认为是多感觉通道信息产生交互的脑区)。这些结果说明了，声音诱发闪光错觉并不全是自动整合加工的结果，还有自上而下的注意参与了视觉和听觉信号的错误整合(Mishra et al., 2010)。此外，Kamke 等人(2012)利用经颅磁刺激(transcranial magnetic stimulation, TMS)研究了声音诱发视觉闪光错觉现象中与注意相关皮层神经网络的作用。研究主要采用不同的视听刺激组合(一个视觉刺激伴随两个听觉刺激，一个视觉刺激，两个听觉刺激)与不同的视听刺激的时间间隔(70 ms 和 160 ms)。结果表明，当 TMS 作用于右角回(angular gyrus, AG)时，被试报告闪光错觉的比例显著降低。因此，他们认为被破坏的角回增强了被试在多感觉事件中的真实知觉,有效地降低了不相关的听觉刺激对视觉事件加工的影响,并且减少了听觉和视觉刺激错误整合的可能性。但是当单一视觉刺激呈现或视听刺激有较长的时间间隔时, TMS 不会改变被试知觉。因此，角回涉及在声音诱发闪光错觉中视觉和声音刺激整合的调节，但是该研究仅限于裂变错觉现象(Kamke et al, 2012)。

近期研究对注意进行了内部区分后发现，内源性注意在多感觉整合中可以以通道注意(Talsma, 2015)的形式影响多感觉整合过程，即对于通道的注意会减弱被试对不被注意的通道内的信息加工，在被注意的通道内的信息加工则得到增强(Spence, 2011)。同时也有研究发现，相比于注意一个特定通道,注意两个通道的成本要更高，这个成本与当注意分配到通道间时在额顶区或颞上皮层中更为强烈的激活有关(Santangelo, Fagioli, & Macaluso, 2010)。近期的一项研究也发现，相对于注意一个通道而言,同时注意两个通道可以影响声音诱发闪光错觉现象，说明了通道分散的注意能够影响声音诱发闪光错觉,但仅限于裂变错觉(于薇，王爱君，张明，2017)。关于通道注意调节感觉主导效应的一个可能的解释为优先进入原则，即同时呈现两个刺激，被试倾向于认为先注意到的刺激先出现。优先进入效应的发生可能是由于被试内源性地注意视觉/听觉通路，也有可能是因为视觉/听觉刺激本身的出现外源性地捕捉了被试的注意(Posner, Nissen, & Klein, 1976; Turatto, Benso, Galfano, & Umiltá, 2002)。因此,当被试将注意以内源性的方式指向视觉通道时，被试会倾向于认为在双通道试次中视觉刺激先于听觉刺激出现；当被试将注意以内源性的方式指向听觉通道时，被试会倾向于认为在双通道试次中听觉刺激先于视觉刺激出现。许多研究采用不同的方式将被试的注意内源性地指向其中一个通道，其中运用最普遍的便是改变不同通道试次出现的比例。如 Sinnett, Spence 和 Soto-faraco (2007)在实验中将单通道视觉目标的比例增加到 60%，与此同时单通道听觉目标的比例减少至 20% (剩下的 20%的试次为双通道试次)来考察内源性的通道注意对

Colavita 视觉主导效应的影响，实验结果发现 Colavita 视觉主导效应变大；而当单通道视觉目标的比例为 20%，单通道听觉目标的比例为 60%时，Colavita 视觉主导效应变小。因此，也说明了这样的操纵可以成功地将被试的注意内源性地转移至出现频次较多的通道上。

本研究在以往研究的基础上，通过将注意内源性地分配到视觉通道或听觉通道，考察内源性的通道注意是否会影响听觉主导效应中声音诱发闪光错觉现象，即通过改变不同通道试次出现的比例，将被试的注意内源性地转移至出现频次较多的通道上，考察内源性地注意视觉通道或听觉通道是否可以影响声音诱发闪光错觉这种听觉主导效应。研究采用经典的声音诱发闪光错觉范式(Shams et al., 2002)，分为 3 个实验，实验 1 为基线实验，即视觉通道试次、听觉通道试次和视听双通道试次的比例为 1:1:1；实验 2 为内源性地注意视觉通道，即视觉通道试次、听觉通道试次和视听双通道试次的比例为 3:1:1；实验 3 为内源性地注意听觉通道，即视觉通道试次、听觉通道试次和视听双通道试次 1:3:1。我们假设内源性地注意视觉/听觉通道可以影响感觉主导效应的方向，因此可以使得声音诱发闪光错觉这种听觉占主导作用的效应产生变化，具体表现为当注意内源性地指向视觉通道时，对视觉通道内信息的神经反应被增强，行为结果表现为被试判断视觉闪光个数的正确率提高，进而表现出闪光错觉效应减小；当注意内源性地指向听觉通道时，对听觉通道内信息的神经反应被增强，行为结果表现为被试判断视觉闪光个数更容易受到听觉声音刺激的干扰，闪光错觉效应增大。

2　实验 1：基线实验

2.1　方法

2.1.1　被试

被试为 23 名大学生(男生 8 人,女生 15 人)，年龄介于 18~26 岁，所有被试视力正常或者矫正视力正常，之前均未参加过类似实验，实验后被试获得一定的报酬。

2.1.2　实验仪器和实验材料

所有刺激呈现在 iiyama MA203DT Vision Master Pro 513 型号的显示器上，屏幕分辨率为 1024×768 像素，刷新率为 85 Hz。实验程序由 Presentation 软件编程(Neurobehavioral Systems Inc)，所有视觉闪光刺激呈现在黑色背景上。其中视觉闪光刺激为呈现在中央注视点下方 5°视角处的白色圆盘(视角为 2°)，呈现的时间为 17 ms。视觉闪光刺激呈现在中央注视点下方 5°视角处是因为，在听觉声音刺激伴随下，视觉闪光刺激位于外周视野时的错觉效应最大(Shams et al., 2002)。实验中的听觉声音刺激通过头戴式铁三角耳机(ATH-WS99)呈现。听觉声音刺激的响度为 75 dB，频率为 3.5 kHz，呈现时间为 7 ms。

2.1.3 实验设计和实验流程

在实验的每个试次中，给被试呈现单个视觉闪光刺激或者连续的两个视觉闪光刺激，在呈现视觉闪光刺激时既可以单独呈现一个视觉闪光刺激，也可以伴随着一个听觉声音刺激或者连续的两个听觉声音刺激。此外，既可以单独呈现连续两个视觉闪光刺激，也可以伴随一个听觉声音刺激或者连续两个听觉声音刺激。除此之外，还可以单独呈现一个听觉声音刺

激或者单独呈现连续的两个声音刺激。因此，实验中所有视觉闪光刺激和听觉声音刺激构成了 8 种实验条件，即 F1、F1B1、F1B2、F2、F2B1、F2B2、B1 和 B2。为了便于论述，这些试次类型采用了上述统一的表述方式。如 F1B2 是指有一个视觉闪光刺激伴随两个听觉声音刺激的试次，而 F2 则是指仅有两个视觉闪光刺激而无听觉声音刺激的试次。F1 和 F2 为"视觉单通道刺激", B1 和 B2 为"听觉单通道刺激"，其余条件为"视听双通道刺激"。实验流程如图 1 所示，第一个听觉声音刺激先于第一个视觉闪光刺激 23 ms 呈现，两个视觉闪光刺激的时间间隔为 50 ms，两个听觉声音刺激的时间间隔为 57 ms。8 种实验条件中可以出现闪光错觉的条件是 F1B2 和 F2B1，即被试针对同样的物理刺激既可以产生错觉，也可以不产生错觉。因此，可以根据被试的反应情况进行事后分类，将上述两种条件分成 F1B2_W (有错觉) vs. F1B2_R (无错觉)以及 F2B1_W (有错觉) vs. F2B1_R (无错觉)。要求被试在整个实验过程中盯住中央注视点来判断他们知觉到一个闪光刺激还是两个闪光刺激，对于听觉单通道刺激并不需要做任何反应。每个被试需要完成 512 个试次，每种实验条件下 64 个试次(F1、F1B1、F1B2、F2、F2B1、F2B2、B1 和 B2)随机呈现。试次间的时间间隔以 100 ms 为步长从 400 ms 到 700 ms 随机。

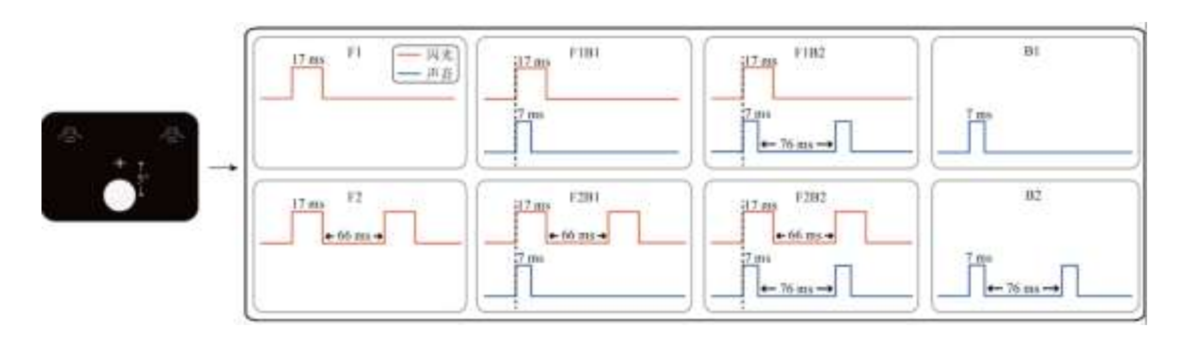


图 1　实验 1 刺激示意图

注：F1 为一个视觉闪光刺激; F2 为两个视觉闪光刺激; B1 为一个听觉声音刺激; B2 为两个听觉声音刺激; F1B1 为一个视觉闪光刺激一个听觉声音刺激; F1B2 为一个视觉闪光刺激两个听觉声音刺激; F2B1 为两个视觉闪光刺激一个听觉声音刺激; F2B2 为两个视觉闪光刺激两个听觉声音刺激。彩图见电子版。

2.2　结果与分析

从表 1 可以看出 F1、F2、F1B1 和 F2B2 条件下被试的正确率都较高，也就是说当视觉刺激单独呈现或者视觉刺激与伴随的听觉刺激个数一致时，被试能够较为准确地做出判断。此外，根据被试的反应情况进行事后分类，将 F1B2 和 F2B1 这两种可产生错觉的条件分成 F1B2_W (有错觉,也称之为裂变错觉) vs. F1B2_R (无错觉,即正确按键)以及 F2B1_W (有错觉,也称之为融合错觉) vs. F2B1_R (无错觉,即正确按键)。通过表 1 可以看出, F1B2 条件的正确率为 34%，显著小于 F1、F1B1、F2 和 F2B2 条件下的正确率($ps < 0.001$), F2B1 条件的正确率为 62%，显著小于 F1、F1B1、F2 和 F2B2 条件下的正确率($ps < 0.001$)。对于视觉刺激与

听觉刺激不匹配的条件(F1B2 条件和 F2B1 条件)，从正确率的角度来看，出现了听觉主导效应，即听觉刺激的个数会影响对视觉刺激个数的判断。

表 1　三个实验各条件下的正确率平均数和标准差(%)

实验	F1	F1B1	F1B2	F2	F2B1	F2B2
实验 1	90 ± 8	94 ± 8	34 ± 24	89 ± 6	62 ± 21	92 ± 9
实验 2	90 ± 6	92 ± 6	50 ± 22	87 ± 7	68 ± 23	96 ± 4
实验 3	90 ± 7	94 ± 6	29 ± 26	88 ± 8	64 ± 25	97 ± 3

注：F1 为一个视觉闪光刺激; F2 为两个视觉闪光刺激; F1B1 为一个视觉闪光刺激一个听觉声音刺激; F1B2 为一个视觉闪光刺激两个听觉声音刺激; F2B1 为两个视觉闪光刺激一个听觉声音刺激; F2B2 为两个视觉闪光刺激两个听觉声音刺激。

3　实验 2：基于视觉通道的内源性注意对声音诱发闪光错觉的影响

3.1　方法

3.1.1　被试

被试为 25 名大学生(男生 9 人，女生 16 人)，年龄介于 18~26 岁，所有被试视力正常或者矫正视力正常，之前均未参加过类似实验，实验后被试获得一定的报酬。

3.1.2　实验仪器和实验材料

实验仪器和实验材料同实验 1。

3.1.3　实验设计和实验流程

实验设计和实验流程同实验 1。不同的是，实验将“视觉单通道刺激”、“听觉单通道刺激”和“视听双通道刺激”的比例设置为 3:1:1。被试的任务是在整个实验过程中都要求盯住中央注视点，判断他们知觉到一个视觉闪光刺激还是两个视觉闪光刺激，针对 B1 和 B2 条件则无需按键反应。每个被试需要完成 768 个试次，其中 F1 和 F2 条件下各 192 个试次,其余各实验条件下各 64 个试次(F1B1、F1B2、F2B1、F2B2、B1 和 B2)，所有刺激随机呈现。试次间的时间间隔以 100 ms 为步长从 400 ms 到 700 ms 随机。

3.2　结果与分析

从表 1 可以看出 F1、F2、F1B1 和 F2B2 条件下被试的正确率都较高，也就是说当视觉刺激单独呈现或者视觉刺激与伴随的听觉刺激个数一致时，被试能够较为准确地做出判断。此外，根据被试的反应情况进行事后分类，将 F1B2 和 F2B1 这两种可产生错觉的条件分成 F1B2_W (有错觉，也称之为裂变错觉) vs. F1B2_R (无错觉，即正确按键)以及 F2B1_W (有错觉，也称之为融合错觉) vs. F2B1_R (无错觉，即正确按键)。通过表 1 可以看出, F1B2 条件的正确率为 50%，显著小于 F1、F1B1、F2 和 F2B2 条件下的正确率($ps < 0.001$), F2B1 条件的正确率为 68%，显著小于 F1、F1B1、F2 和 F2B2 条件下的正确率($ps < 0.001$)。对于视觉刺激与听觉刺激不匹配的条件(F1B2 条件和 F2B1 条件)，从正确率的角度来看,出现了听觉主导效应，即听觉刺激的个数会影响对视觉刺激个数的判断。

4　实验 3：基于听觉通道的内源性注意对声音诱发闪光错觉的影响

4.1 方法

4.1.1 被试

被试为 24 名大学生(男生 9 人，女生 15 人)，年龄介于 18~26 岁，所有被试视力正常或者矫正视力正常，之前均未参加过类似实验，实验后被试获得一定的报酬。

4.1.2 实验仪器和实验材料

实验仪器和实验材料同实验 1。

4.1.3 实验设计和实验流程

实验设计和实验流程同实验 1。不同的是，实验将“视觉单通道刺激”、“听觉单通道刺激”和“视听双通道刺激”的比例设置为 1:3:1。被试的任务是在整个实验过程中都要求盯住中央注视点，判断他们知觉到一个视觉闪光刺激还是两个视觉闪光刺激，针对 B1 和 B2 条件则无需按键反应。每个被试需要完成 768 个试次，其中 B1 和 B2 条件下各 192 个试次,其余各实验条件下各 64 个试次(F1、F1B1、F1B2、F2、F2B1 和 F2B2)，所有刺激随机呈现。试次间的时间间隔以 100 ms 为步长从 400 ms 到 700 ms 随机。

4.2 结果与分析

从表 1 可以看出 F1、F2、F1B1 和 F2B2 条件下被试的正确率都较高，也就是说当视觉刺激单独呈现或者视觉刺激与伴随的听觉刺激个数一致时，被试能够较为准确地做出判断。此外，根据被试的反应情况进行事后分类，将 F1B2 和 F2B1 这两种可产生错觉的条件分成 F1B2_W (有错觉，也称之为裂变错觉) vs. F1B2_R (无错觉，即正确按键)以及 F2B1_W (有错觉，也称之为融合错觉) vs. F2B1_R (无错觉，即正确按键)。通过表 1 可以看出, F1B2 条件的正确率为 29%，显著小于 F1、F1B1、F2 和 F2B2 条件下的正确率($ps < 0.001$), F2B1 条件的正确率为 64%，显著小于 F1、F1B1、F2 和 F2B2 条件下的正确率($ps < 0.001$)。对于视觉刺激与听觉刺激不匹配的条件(F1B2 条件和 F2B1 条件)，从正确率的角度来看，出现了听觉主导效应，即听觉刺激的个数会影响对视觉刺激个数的判断。

5 结果综合分析

5.1 正确率

将实验 1、实验 2 和实验 3 中视觉闪光刺激与听觉声音刺激个数不匹配试次(F1B2 条件和 F2B1 条件)的正确率进行 3 (实验 1 vs. 实验 2 vs. 实验 3) × 2 (F1B2 vs. F2B1)的重复测量方差分析。结果发现，不同实验的主效应显著,F (1, 22) = 4.01, $p < 0.05$, $\eta^2 = 0.16$；错觉条件的主效应显著，F (1, 22) = 60.73, $p < 0.001$, $\eta^2 = 0.73$；两者的交互作用不显著，F (1, 22) = 1.16, $p > 0.05$。为了进一步考察 F1B2 和 F2B1 条件下各个实验的潜在差异，将 F1B2 和 F2B1 条件下各个实验进行单因素的方差分析，考察内源性注意视觉通道和内源性注意听觉通道对声音诱发闪光错觉中裂变错觉和融合错觉的影响(图 2)。结果发现，在 F1B2 条件下，不同实验间差异显著，F (1, 2) = 5.11, $p < 0.01$, $\eta^2 = 0.23$。事后检验结果显示，实验 1 和实验 2 之间差异显著($p < 0.05$)；实验 2 和实验 3 之间差异显著($p < 0.005$)；而实验 1 和实验 3 之间差异不显著

($p > 0.05$)。在 F2B1 条件下，不同实验间差异不显著($ps > 0.05$)。因此，从结果上可以看出，基于通道的内源性注意可以影响声音诱发闪光错觉中的裂变错觉(F1B2)，具体表现在，基于视觉通道的内源性注意条件下 F1B2 的正确率显著增加，而基于听觉通道的内源性注意条件下 F1B2 的正确率减少，但不显著。内源性注意视觉通道和内源性注意听觉通道间 F1B2 的正确率存在显著差异。而对于声音诱发闪光错觉中的融合错觉(F2B1)而言，其并不受内源性的通道注意所影响。

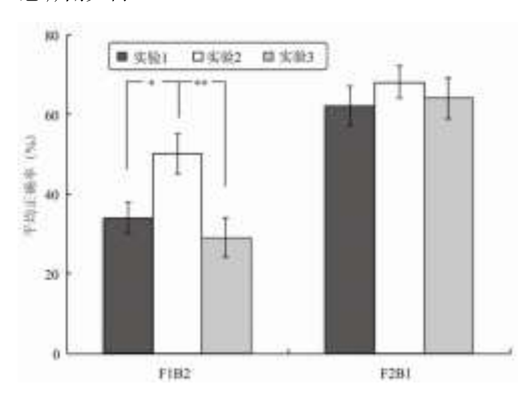


图 2　F1B2 和 F2B1 的正确率在三种实验间的差异比较

注：F1B2 为一个视觉闪光刺激两个听觉声音刺激; F2B1 为两个视觉闪光刺激一个听觉声音刺激。

尽管裂变错觉和融合错觉是一个稳定且普遍的现象，但每个人的错觉量存在差异。因此，为了避免这种个体差异，Andersen 等人(2004)采用了一个新的指标来表示裂变错觉和融合错觉效应量的大小，这个指标即比值比率。求比值比率，首先要将所有被试在双通道条件下能产生裂变错觉的试次总数求出，接着求出其中真正产生裂变(即感觉到的闪光个数多于实际呈现的闪光个数)的试次数，二者的差值就是没有产生裂变错觉的试次数，用产生裂变错觉的试次总数与没有产生裂变错觉的试次数量进行比较，可以得到一个比值。同理，求出单通道条件下相应的比值(单通道条件下能产生裂变错觉的试次即闪光个数与双通道条件下闪光个数一致的试次)。将这两个比值再进行比较，得到的比值比率能更清晰地反映出裂变错觉的效应量大小。当比值比率大于 1 的时候，说明在双通道条件下产生的裂变错觉比单通道条件下更大，即听觉声音影响了对视觉闪光的判断，比值比率的值越大，说明错觉效应量越大。融合错觉的计算与此一致(Andersen et al., 2004)。按照 Andersen 等人(2004)的算法计算了裂变错觉和融合错觉的量,也能够更加清晰地比较出不同实验条件下错觉效应的大小(表 2 和表 3)。从表 2 和表 3 中可以看出，无论是在实验 1、实验 2 还是在实验 3 中，相对于融合错觉而言,裂变错觉的效应都更大。此外，也能够反映出 F1B2 条件下，实验 1 和实验 2 存在显著差异，实验 2 和实验 3 存在显著差异，但实验 1 和实验 3 之间并不存在显著差异；F2B1 条件下，三个实验之间并不存在显著差异。

表 2　各实验中裂变错觉的比值比率

刺激类型	实验 1			实验 2			实验 3		
	裂变错觉	无裂变	总数	裂变错觉	无裂变	总数	裂变错觉	无裂变	总数
视听刺激	972	500	1472	800	800	1600	1091	445	1536
视觉刺激	147	1325	1472	480	4320	4800	154	1382	1536
比值比率		17.5			9			21.95	

表 3　各实验中融合错觉的比值比率

刺激类型	实验 1			实验 2			实验 3		
	融合错觉	无裂变	总数	融合错觉	无裂变	总数	融合错觉	无裂变	总数
视听刺激	559	913	1472	512	1088	1600	553	983	1536
视觉刺激	162	1310	1472	624	4176	4800	184	1352	1536
比值比率		4.95			3.15			4.13	

5.2　反应时

将 F1B2 和 F2B1 条件下不同实验间的反应时进行 3 (实验：实验 1 vs. 实验 2 vs. 实验 3) × 4 (条件：F1B2_R vs. F1B2_W vs. F2B1_R vs. F2B1_W) 的混合重复测量方差分析(见表 4)。结果发现，实验的主效应显著, F (1, 2) = 23.91, p < 0.001, η^2 = 0.41；条件的主效应显著, F (3, 69) = 124.81, p < 0.001, η^2 = 0.64；实验与条件的交互作用显著, F (6, 69) = 2.47, p < 0.05, η^2 = 0.07。为了进一步了解不同实验和裂变错觉、融合错觉之间潜在的交互关系，我们分别将不同实验下的 F1B2_R 和 F1B2_W 以及 F2B1_R 和 F2B1_W 进行比较分析。结果发现，实验 1 中, F1B2_R 条件下的反应时(889 ms)显著长于 F1B2_W 条件下的反应时(839 ms), t (22) = 4.07, p = 0.001, d = 1.22；实验 2 中, F1B2_R 条件下的反应时(798 ms)显著长于 F1B2_W 条件下的反应时(710 ms), t (24) = 5.44, p < 0.001, d = 1.56；实验 3 中, F1B2_R 条件下的反应时(732 ms)显著长于 F1B2_W 条件下的反应时(707 ms), t (23) = 2.15, p < 0.05, d = 0.63。但是在实验 1、实验 2 以及实验 3 中，F2B1_R 条件下的反应时和 F2B1_W 条件下的反应时之间均无显著差异，具体结果为，实验 1 中, F2B1_R 条件下的反应时(757 ms)和 F1B2_W 条件下的反应时(762 ms)间无显著差异, t < 1；实验 2 中, F2B1_R 条件下的反应时(640 ms)和 F1B2_W 条件下的反应时间无显著差异(639 ms), t < 1；实验 3 中, F2B1_R 条件下的反应时(581 ms)和 F1B2_W 条件下的反应时无显著差异(613 ms), t (23) = 1.99, p > 0.05。

表 4　各实验 F1B2 和 F2B1 条件下的反应时平均数和标准差(ms)

实验	F1B2_R	F1B2_W	F2B1_R	F2B1_W
实验 1	889 ± 90	839 ± 70	757 ± 83	762 ± 92
实验 2	798 ± 95	710 ± 108	640 ± 104	639 ± 88
实验 3	732 ± 89	707 ± 76	581 ± 86	613 ± 99

注：F1B2_R 为一个视觉闪光刺激两个听觉声音刺激，并未产生错觉条件；F1B2_W 为一个视觉闪光刺激两个听觉声音刺激，产生错觉条件；F2B1_R 为两个视觉闪光刺激一个听觉声音刺激，并未产生错觉条件；F2B1_W 为两个视觉闪光刺激一个听觉声音刺激，产生错觉条件。

上述分析的结果表明，在 F1B2 条件中(裂变错觉)，相对于有错觉条件(F1B2_W)，无错觉条件下(F1B2_R)的反应时都明显增长，即被试需要付出更多注意资源才能正确辨别视觉闪光刺激的个数，这种现象无论在实验 1 (50 ms)、实验 2 (88 ms)还是实验 3 (26 ms)都存在。将三个实验中 F1B2_R 和 F1B2_W 的差值进行单因素方差分析，结果发现显著性差异，F (2, 68) = 4.61, p < 0.05, η^2 = 0.22。进一步事后检验发现，实验 1 和实验 2 之间差异不显著(p > 0.05)，实验 1 和实验 3 之间差异不显著(p > 0.05)，但实验 2 和实验 3 之间差异显著(p < 0.005)。研究结果说明，不管是否存在注意的影响，以及内源性注意视觉通道还是听觉通道，裂变错觉这种听觉主导效应的现象均稳定存在。然而，在 F2B1 条件中(融合错觉)，有错觉条件(F2B1_W)与无错觉条件下(F2B1_R)的反应时之间并无显著差异，这种现象在三个实验同样存在，说明了不管是否存在注意的影响，以及内源性注意视觉通道还是听觉通道，融合错觉这种听觉主导效应的现象在反应时水平上并未得到体现。

此外，为了比较不同的实验之间 F1B2_W 和 F2B1_W 的差异，即不同内源性注意操纵条件下，裂变效应和融合效应相应的变化，我们将 F1B2_W 和 F2B1_W 条件下不同实验间的反应时分别进行了单因素方差分析(实验 1 vs. 实验 2 vs. 实验 3)。结果发现，在 F1B2_W 条件下，不同实验间的反应时差异显著，F (2, 69) = 17.49, p < 0.001, η^2 = 0.37。事后检验结果显示，实验 1 和实验 2 之间差异显著(p < 0.001)；实验 1 和实验 3 之间差异显著(p < 0.001)；而实验 2 和实验 3 之间差异不显著(p > 0.05)。在 F2B1_W 条件下，不同实验间的反应时差异显著 F (2, 69) = 16.94, p < 0.001, η^2 = 0.35。事后检验结果显示，实验 1 和实验 2 之间差异显著(p < 0.001)；实验 1 和实验 3 之间差异显著(p < 0.001)；而实验 2 和实验 3 之间差异不显著(p > 0.05)。

6　讨论

研究采用经典的声音诱发闪光错觉范式(Shams et al., 2000; 2002)，通过将注意内源性地分配到视觉通道或听觉通道，即通过改变不同通道试次出现的比例，将被试的注意内源性地转移至出现频次较多的通道上，考察内源性地注意视觉通道或听觉通道是否可以影响声音诱发闪光错觉这种听觉主导效应。研究结果发现，存在经典的声音诱发闪光错觉现象,并且裂变错觉的效应大于融合错觉的效应。此外，研究还发现，当注意内源性地指向不同感觉通道时会影响裂变错觉的大小，但并不影响融合错觉的大小。具体表现为，当注意内源性地指向视觉通道时,相对于基线条件，裂变错觉的量显著减小(图 2)；当注意内源性地指向听觉通道时，相对于基线条件，裂变错觉的量增大，但与基线条件差异不显著(图 2)。本研究之所以将比例操纵作为被试间变量的原因有：其一，由于目前的实验时间已经较长，加之实验环境为低亮度的屏蔽室内，实验材料的呈现方式是白色闪光呈现在黑色背景上，刺激的对比很强，如果将 3 个实验整合为一个实验，实验的时间比较长，被试会产生疲劳效应；其二，声音诱

发闪光错觉是一种非常稳定的听觉主导效应，虽然在声音诱发闪光错觉这一现象中存在个体差异，但是从总体上来看，实验中每个被试都存在这种错觉效应，尽管有的被试错觉量较大，有的较小。此外，对于实验 1，我们前期用多批被试做过多次重复实验，其结果基本一致(也见于薇等, 2017)。因此，本研究将比例调控设置为被试间变量并不会受到被试的差异而对实验结果产生影响。其三，在 Colavita 视觉主导效应的研究中, Sinnett 等人(2007)也是在不同的实验中(被试不同)，设置了不同通道刺激的比例，研究结果发现了通过比例设置操纵的内源性通道注意可以影响 Colavita 视觉主导效应的大小。

从正确率的结果上来看，本研究 3 个实验均发现经典的声音诱发闪光错觉听觉主导效应，并且裂变错觉效应大于融合错觉，这与以往研究结果一致(Shams et al., 2000, 2002; Shams, Ma, & Beierholm, 2005; Watkins, Shams, Tanaka, Haynes, & Rees, 2006; Watkins, Shams, Josephs, & Rees, 2007; Wozny et al., 2008; 于薇等, 2017)。除此之外，本研究发现裂变错觉会受到内源性通道注意的影响，但融合错觉则不会受内源性通道注意的影响(图 2)。原因在于，首先，以往研究发现，相对于裂变错觉，融合错觉不稳定(Shams et al., 2000; Wozny et al., 2008)，表现为裂变错觉这种错觉效应尽管改变诸如形状、对比度、大小、视觉刺激呈现的时间、频率和强度、听觉刺激的呈现时间，以及视觉刺激和听觉刺激在时间上和空间上的相对位置等条件仍然稳定存在(Shams et al., 2000, 2002; Shams et al., 2005; Watkins et al., 2006; Watkins et al., 2007)。其次，裂变错觉会受到注意资源的影响而产生错觉量大小的变化，但是融合错觉并不受到注意资源的影响。具体表现为，相对于注意只集中在视觉单通道条件下，当注意既分配到视觉通道也分配到听觉通道时，裂变错觉量减小，而融合错觉不受其影响(于薇等，2017)。以往研究表明,注意除能指向某一空间位置外，还能指向某一感觉通道(Talsma，2015)，即对被注意的通道内的信息加工得到增强，而对不被注意的通道内的信息加工则得到减弱(顾吉有，吕勇, 2016; Spence, 2011; 孙远路 等, 2011)。因此，当注意指向一个感觉通道时，对此通道内信息的神经反应被增强，而对被忽视通道内信息的神经反应则被抑制。声音诱发闪光错觉现象是多感觉整合中的一种听觉主导现象，涉及到视觉和听觉两个通道，本研究中，将注意内源性地指向视觉通道时，对视觉通道内信息的神经反应被增强，行为结果表现为被试判断视觉闪光个数的正确率提高,进而表现出裂变错觉效应减小(图 2 和表 2)；而将注意内源性地指向听觉通道时，对听觉通道内信息的神经反应被增强，行为结果表现为被试判断视觉闪光个数更容易受到听觉声音刺激的干扰，使得被试反应的正确率下降，进而表现出裂变错觉效应增大的趋势(图 2 和表 2)。

本研究通过比例设置的方式将注意内源性地指向视觉通道和听觉通道，结果也发现，同 Colavita 视觉主导效应一样，声音诱发双闪光错觉(裂变错觉)也会受到内源性通道注意的影响，使得裂变错觉的效应减小或增大。研究结果与 Sinnett 等人(2007)考察内源性的通道注意对 Colavita 视觉主导效应影响的结果一致，发现了通过比例设置操纵的内源性通道注意可以影响感觉主导效应的大小。实验 2 中注意内源性地指向视觉通道，实验 3 中注意内源性地指向听觉通道。这里我们强调的是不同感觉通道的注意，研究中实验任务为判断视觉通道刺激

的个数，因此，当注意内源性地指向视觉通道时，听觉通道的刺激被试并不需要注意，因此可以视为集中注意视觉单一通道。但是当注意内源性地指向听觉通道时，被试一方面需要注意听觉通道的刺激，另一方面也需要注意视觉通道的刺激(任务需求),因此可以视为分散注意视觉通道和听觉通道。有研究发现，相比于注意一个特定通道，注意两个通道的成本要更高，这个成本与当注意分配到通道间时额顶区或颞上皮层中更为强烈的激活有关(Degerman et al., 2007; Santangelo et al., 2010)。我们近期的一项研究也发现，相对于注意一个通道而言，同时注意两个通道可以影响声音诱发闪光错觉现象，说明了通道分散的注意能够影响声音诱发闪光错觉(于薇等，2017)。关于通道注意调节感觉主导效应的一个可能的解释为优先进入原则，即同时呈现两个刺激，先注意到的刺激被试倾向于认为其先出现。优先进入效应的发生可能是由于被试内源性地注意视觉/听觉通路，也有可能是因为视觉/听觉刺激本身的出现外源性地捕捉了被试的注意(Posner et al., 1976; Turatto et al., 2002)。因此，当被试将注意以内源性的方式指向视觉通道时，被试会倾向于选择性地将注意集中在视觉通道,导致对视觉通道信息的神经反应增强，即被试在判断视觉闪光刺激的个数时不容易受到听觉通道刺激的干扰；当被试将注意以内源性的方式指向听觉通道时，被试需要将注意分散在视觉通道和听觉通道，相比于集中注意单一视觉通道，被试对听觉通道信息的神经反应增强,即被试在判断视觉闪光刺激的个数时更容易受到听觉通道刺激的干扰。

本研究中，注意内源性地指向视觉通道可以视为被试在判断视觉闪光刺激的个数时不容易受到听觉通道刺激的干扰；注意内源性地指向听觉通道可以视为被试在判断视觉闪光刺激的个数时更容易受到听觉通道刺激的干扰。在多感觉整合/竞争的研究中，一个重要的影响因素就是分心刺激在多大程度上会和目标刺激竞争被试的注意资源。根据 Lavie (2005)的知觉负载理论，相比于无负载或者低负载条件，被试在高负载的条件下需要将注意全部集中在当前的任务，因此，如果被试的注意被分散到多个感觉通道间(高负载条件)，那么被试在其中某一个通道条件下分配的注意资源较少，表现为行为反应较差(Tang，Wu，& Shen, 2016)。因此，实验 2 被试只需注意单一的视觉通道，被试在完成视觉判断任务时并不受到听觉通道刺激的干扰，表现为错觉量的减小；实验 3 被试需要注意视觉和听觉双通道，被试在完成视觉判断任务时会受到听觉通道刺激的干扰，表现为错觉量出现增大的趋势。

7 结论

基于通道的内源性注意可以影响声音诱发双闪光错觉(裂变错觉)听觉主导效应，即内源性地注意视觉通道使得裂变错觉减小，内源性地注意听觉通道使得裂变错觉表现出增大的趋势。

参考文献

Abadi, R. V., & Murphy, J. S. (2014). Phenomenology of the sound-induced flash illusion. *Experimental Brain Research*, *232*(7), 2207–2220.

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research, 21*(3), 301–308.

Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology, 25*(2), 231–235.

Degerman, A., Rinne, T., Pekkola, J., Autti, T., Jaaskelainen, I.P., Sams, M., & Alho, K. (2007). Human brain activity associated with audiovisual perception and attention. *Neurolmage, 34*(4), 1683–1691.

Gu, J. Y., & Lv, Y. (2016). The different effects of selective attention and divided attention on multisensory integration. *Studies of Psychology and Behavior, 14*(2), 202–206.

[顾吉有，吕勇. (2016). 选择性注意和分配性注意对多感觉整合的不同影响. *心理与行为研究, 14*(2), 202–206.]

Kamke, M. R., Vieth, H. E., Cottrell, D., & Mattingley, J. B. (2012). Parietal disruption alters audiovisual binding in the sound-induced flash illusion. *Neurolmage, 62*(3), 1334–1341.

Kumpik, D. P., Roberts, H. E., King, A. J., & Bizley, J. K. (2014). Visual sensitivity is a stronger determinant of illusory processes than auditory cue parameters in the sound-induced flash illusion. *Journal of Vision, 14*(7), 60-65.

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences, 9*(2), 75–82.

Mishra, J., Martínez, A., & Hillyard, S. A. (2010). Effect of attention on early cortical processes associated with the sound-induced extra flash illusion. *Journal of Cognitive Neuroscience, 22*(8), 1714–1729.

Mishra, J., Martínez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience, 27*(15), 4120–4131.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83*(2), 157–171.

Santangelo, V., Fagioli, S., & Macaluso, E. (2010). The costs of monitoring simultaneously two sensory modalities decrease when dividing attention in space. *Neurolmage, 49*(3), 2717–2727.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature, 408*(6814), 788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*(1), 147–152.

Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport, 16*(17), 1923–1927.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Attention, Perception, & Psychophysics, 69*(5), 673–686.

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention, Perception, & Psychophysics, 73*(4), 971–995.

Sun, Y-L., Hu, Z-H., Zhang, R-L., Xun, M-M., Liu, Q., & Zhang, Q-L. (2011). An investigation on the effect factors in the paradigm of multisensory integration. *Acta Psychologica Sinica, 43* (11), 1239–1246.

[孙远路，胡中华，张瑞玲，寻茫茫，刘强，张庆林. (2011). 多感觉整合测量范式中存在的影响因素探讨. *心理学报, 43* (11), 1239–1246.]

Talsma, D. (2015). Predictive coding and multisensory integration: An attentional account of the multisensory mind. *Frontiers in Integrative Neuroscience, 9*, 19.

Tang, X. Y., Wu, J. L., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience & Biobehavioral Reviews, 61*, 208–224.

Turatto, M., Benso, F., Galfano, G., & Umiltá, C. (2002). Nonspatial attentional shifts between audition and vision. *Journal of Experimental Psychology: Human Perception & Performance, 28*(3), 628–639.

Van Erp, J. B. F. V., Philippi, T. G., & Werkhoven, P. (2013). Observers can reliably identify illusory flashes in the illusory flash paradigm. *Experimental Brain Research, 226*(1), 73–79.

Watkins, S., Shams, L., Josephs, O., & Rees, G. (2007). Activity in human V1 follows multisensory perception. *NeuroImage, 37*(2), 572–578.

Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage, 31*(3), 1247–1256.

Wozny, D. R., Beierholm, U. R., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of Vision, 8*(3), 1–11.

Whittingham, K. M., Mcdonald, J. S., & Clifford, C. W. G. (2014). Synesthetes show normal sound-induced flash fission and fusion illusions. *Vision Research, 105*, 1–9.

Yu, W., Wang, A. J., & Zhang, M. (2017). Effect of selective and divided attentions on auditory dominance in multisensory integration. *Acta Psychologica Sinica, 49*(2), 164–173.

[于薇，王爱君，张明. (2017). 集中和分散注意对多感觉整合中听觉主导效应的影响. *心理学报, 49*(2), 164–173.]

The Effects of modal-based endogenous attention on sound-induced flash illusion

ZHANG Ming[1]; TANG Xiaoyu[2]; YU Wei[3]; NING Bo[1]; WANG Zhinan[1]; WANG Aijun[1];

([1] *Department of Psychology, Research Center for Psychology and Behavioral Sciences, Soochow University, Suzhou* 215000, *China*)

([2] *School of Psychology, Liaoning Collaborative Innovation Center of Children and Adolescents Healthy Personality Assessment and Cultivation, Liaoning Normal University, Dalian* 116029, *China*)

([3] *Admission and Employment Office, Changchun University of Chinese Medicine, Changchun* 130117, *China*)

Abstract

The sound-induced flash illusion (SIFI) is an auditory-dominant phenomenon in multisensory integration in which the perception of brief visual flashes may be qualitatively altered by concurrent brief sounds. It has been well documented how attention modulates this illusion. It remains unknown, however, how modal-based endogenous attention interferes with the SIFI. In the present study, we adopted the classical paradigm of the SIFI and directed the endogenous attention to the visual modal or the auditory modal to investigate the effect of modal-based endogenous attention on SIFI.

Experiment 1 asked the participants to conduct the classical task of the SIFI as the baseline. In experiment 2, the ratio of the visual modal, the auditory modal, and the audio-visual modal was 3:1:1; that is, the endogenous attention was directed to the visual modal. In experiment 3, the ratio of the visual modal, auditory modal and audio-visual modal was 1:3:1; that is, the endogenous attention was directed to the auditory modal. In the present study, we asked the participants to judge the number of flashes, and we focused mainly on the occurrence of fission illusions (when a single visual flash is accompanied by two auditory bleeps and is perceived incorrectly as two flashes) and fusion illusions (when two visual flashes are accompanied by a single bleep and are perceived incorrectly as a single flash). The hypothesis was that directing endogenous attention to the visual or auditory modal could affect the auditory dominance.

A repeated measures ANOVA was analyzed for fission and fusion illusions separately, and the results showed that there were significant differences among the three experiments, F (1, 2) = 5.11, p < 0.01, η^2 = 0.23. The occurrence of fission illusions was decreased when endogenous attention was directed to the visual modal (50%) and was increased when endogenous attention was directed to the auditory modal (71%). However, regardless of whether endogenous attention was directed to the auditory or visual modal, fusion illusion was not affected by attention (ps > 0.05). In addition, we also pooled the response counts across all of the participants and for both fission and fusion stimuli. The results showed that the odds ratio for fission was greater than that for fusion, and there was a difference among the three experiments for fission (17.5 vs. 9 vs. 21.95), and no effect for fusion (4.95 vs. 3.15 vs. 4.13).

The results indicated that modal-based endogenous attention can affect the occurrence of fission illusions but does not affect the occurrence of fusion illusions. The present study provides insight into the top-down factors that can modulate sound-induced flash illusions.

Key words: endogenous attention; auditory dominance; sound-induced flash illusion; visual and auditory modals